Knowledge Extraction for the Product Development Process Based on Ontology-Driven Semantic Interoperability

Athon F. C. S. de Moura Leite, Matheus B. Canciglieri, Anderson L. Szejka, Osiris Canciglieri Junior, and Robert I. M. Young

Abstract The current product development scenario challenges manufacturing industry to deliver improved products to the market while ensuring improved quality. To ensure the best value, companies need to share product requirements effectively from various sources and domains, but there are still misinterpretation and mistakes on this process, regarding semantic interoperability obstacles in the context of the process of requirements' gathering, translating, and reusing them. To help in solving these problems, this study proposes an approach to aid the gathering of product knowledge, extracting it, and translating the knowledge for further use along the Integrated Product Development Process (IPDP). The approach consisted of analysing current issues of the topics, followed by the development of a novel approach, to then be further tested in an experimental case. Issues found on literature point to research gaps related to semantic reconciliation and the extraction of knowledge perspectives, in which semantic issues are approached through different points of view and multiple domains. The proposed approach considers unprocessed product requirements, further translated in features, and by that refining product knowledge during IPDP and enabling it to be reusable. The proposed solution shows a new method to collect and translate product requirements, while gathering its knowledge and transforming it in product features. The tests in an experimental case have shown a

A. F. C. S. de Moura Leite · M. B. Canciglieri · A. L. Szejka (✉) · O. Canciglieri Junior
Industrial and Systems Engineering Graduate Program (PPGEPS), Pontifical Catholic University of Paraná (PUCPR), Curitiba, Paraná, Brazil
e-mail: anderson.szejka@pucpr.br

A. F. C. S. de Moura Leite
e-mail: athon.leite@pucpr.edu.br

M. B. Canciglieri
e-mail: matheus.canciglieri@pucpr.edu.br

O. Canciglieri Junior
e-mail: osiris.canciglieri@pucpr.br

R. I. M. Young
School of Mechanical, Electrical and Manufacturing Engineering, Loughborough University, Loughborough, Leicestershire, UK
e-mail: r.i.young@lboro.ac.uk

B. Archimède et al. (eds.), *Enterprise Interoperability IX*, Proceedings of the I-ESA Conferences 10, https://doi.org/10.1007/978-3-030-90387-9_3

reduction in development time, and an increase in product quality, having significant impacts in reducing costs of development while ensuring correct communication and effectively sharing information.

Keywords Semantic web approaches · Knowledge management · Reference ontologies · Standardisation management and strategies

1 Introduction

Current product development is marked by complex requirements, higher standards of quality, and products that constantly need to adapt to fulfil customer's needs. This dynamism is tied to the trends on integrated manufacturing systems and Industry 4.0. The latter accompanies the use of technologies and methods to improve information sharing considering multiple domains of knowledge [1]. In this context, an Integrated Product Development Process (IPDP) deals with multiple domains of knowledge as a way to gather requirements to product development. The product requirements need to be well defined and shared with little or no loss of meaning during IPDP, in order to avoid misinterpretation, incoherency, and other issues during development, as increased costs and delays [2].

The contemporary practise is still not coping with those issues, that are semantic in nature, within a multiple domain environments, as product requirements must be consistent, clear, stand alone, measurable, testable, unique, unambiguous, and verifiable [1]. A significant portion of the issues in the requirements are related to them having different taxonomies for their representation, different points of view from agents during development, and limitations regarding the process of translating the knowledge in product development requirements. As a result, the misinterpretation of product requirements is related directly to wrong assumptions based on different information from heterogeneous domains [3].

In order to address the aforementioned issues, semantic interoperability (SI) has as its objective the effective information sharing in collaborative environments based on heterogenous domains. SI is being applied in the domains product design and manufacturing, in order to reduce semantic issues and cope with different sources of information [3]. Problems still persist, though, regarding their implementation, more specifically within the methods to extract and translate product requirements from heterogeneous sources of information, as well as standardising them [3].

This research has as its main objective the development of an approach to solve those issues by gathering, organising, and translating standardised product requirements and by that reducing the heterogeneity of interpretation during the IPDPD within a heterogeneous domain environment.

2 Conceptual Background

2.1 Integrated Product Development Process

Current IPDP relies on heterogeneous domains of knowledge, involving agents from different backgrounds and varied experience. Authors [4–7] cite ideal development cycles as integrated, collaborative, and interoperable, since as long as the product information is well defined, the misinterpretation of information and semantic barriers occurrence will be reduced [7].

Current models emphasise the systematisation process of IPDP, which was originally depicted as linear, with subsequent activities starting right after the conclusion of their predecessor [3]. Current approaches redefine that linear structure by inserting the notion of parallel activities across the product development [5, 7].

Research points to the necessity of ensuring interoperability in product and manufacturing in IPDP, as misinterpretation issues happen while product development requires multiple knowledge domains [4]. Research found in [8–10] presented the potential to use ontological methods to formalise knowledge in product and/or manufacturing models.

2.2 Ontology-Driven Semantic Interoperability

The use of ontologies has increased the development of shared representations. Recent research, as depicted in [9, 10], shows that the ability for sharing semantics across product and manufacturing representations can be supported by ontological formalisms. Ontologies are recognised as an important technology to cope with semantic interoperation issues [11]. Its formal structure provides machine-processed semantics of varied knowledge sources [12].

Despite their contributions, even when ontology-based methods are used, in order to assure shared semantics, semantic heterogeneity and their related issues are still unavoidable. Because of that, methods for proper ontology mapping are being developed to improve the semantics between ontologies representing domains that need interoperation [13].

Ontologies may be categorised in three distinct levels of abstraction during their application, depending on their aim [9]:

- Foundation Ontology, which is an ontology that is suited for general concepts and relationships, usable in heterogeneous domains;
- Reference Ontology, which is domain-specific ontology, being reusable in the same domain to perform different tasks; and
- Application Level Ontology, which represents knowledge that is specific and dedicated to unique tasks.

Fig. 1 Methodological procedures

3 Materials and Methods

This research uses as methodological procedures a qualitative literature review and an experimental case. Firstly, a literature review addressing the main issue and its dimensions is proposed, to identify the knowledge gap in which the solution will be developed. Later phases regard the creation and explanation of the approach steps, as depicted in Fig. 1.

The approach will be tested through an experimental case in which a real product will be scrutinised and compared considering its real development metrics and the ones obtained with the implementation of the proposed approach.

4 Literature Review

4.1 Cross-Domain Issues

In past decades, a few models for representation of standardised information structure were developed in heterogeneous domains. For instance, Unified Modelling Language (UML), Domain-Specific Language (DSL), and others [14, 15]. While presenting a way to formalise knowledge representation in different domains in a standardised manner, these models are not able to cope with the dynamic nature of product's requirements and knowledge from different phases of IPDP in a semantically accurate way [16].

Recently, a few models are considering the consistency of requirements and performance in environments with dynamic requirements. Authors in [17] explain a framework to support the design of cyber-physical systems, using design rationale and linking various system parameters and requirements coming from different sources. As shown in [18], authors investigate the manufacturing domain and the process of requirement gathering in different domains. The research in [17] and [18] combined different models to obtain verifiable and valid information within the context of dynamic requirements. However, the final word from specialists is still used remarkably, while translating these requirements in both models. This praxis might result in issues that are semantic in nature, as there might be significant subjectivity in the methods in which each specialist decides, due to their different comprehension of a domain.

4.2 Cross-IPDP Phase Issues

Communication in IPDP is based on the semantic interpretations of each agent [19]. In different phases of IPDP, heterogeneous sets of information might cause misinterpretation due to different meanings for a single term. That is a result of the different background of agents (e.g. product design, engineering, etc.) and their different levels of experience in the stages of product development [16, 17]. Authors in [16] state that knowledge that is required for a single product development stage might have different impacts in later activities, due to the dynamic nature of IPDP and the heterogeneity among their agents.

Currently, as shown in [18], research proposes a formalisation through semantic annotations for applications to semantically interoperate. However, there are no annotations that represent dynamic requirements, as well as automated ways of extracting them. In [20], the author presents an approach to solve cross-IPDP issues based on an ontology that is model driven, but exclusively to crossing two domains.

4.3 Cross-Requirement Issues

Requirements represent the main input in an ontology-driven semantically interoperable system related to IPDP. Their representation needs to be "semantically whole", in order to avoid negative issues, by using clear and well-defined axioms and statements [21]. Despite that, the poor abstraction of statements, in most cases, ends up generating interpretations that are divergent. This results in negative effects related to comprehension, uniqueness, and, in a significant portion, traceability of information.

In semantic interoperation, the comprehensible, unique, and traceable information is able to prevent inconsistencies in the product development and its manufacturing. In [21], a framework to cope with the issues and enable semantic interoperability is presented, in accordance with the previous statement. However, this framework does not ensure the traceability of the requirements and no optimisation of the process and structuration method of knowledge gathering. In [22], the authors presented a model that considers multiple domains, ensuring the traceability of information through verification and validation methods, but limited only to early phases of the development of a product. Current research, as shown in [1, 18, 21], shows the necessity to standardised procedures to extract information and knowledge, ensuring traceability through validation and verification, however, not considering the extraction of requirements. In [1] and [23], authors point out that future interoperable representations must consider knowledge extraction to ensure standardised knowledge gathering. In this sense, the proposed knowledge gap relates to "an automated product knowledge extraction in a multi-domain and interoperable environment that standardises knowledge in a holistic approach to IPDP, avoiding semantic issues".

5 Approach

In order to aid in the process of filling the exposed knowledge gap, an approach to increase automation in the process of gathering, extracting, and translating product knowledge into reference ontologies is proposed in Fig. 2, using IDEF0 notation. Such approach has as its goal the application of concepts and tools of semantic interoperability, in order to develop an interoperable environment. The concept is that this environment is able to represent and further translate knowledge among different phases of IPDP, analysing its consistency and reducing the negative effects caused by heterogeneous knowledge sources.

The approach is an extension of the Interoperable Product Design and Manufacturing System (IPDMS) model proposed in [1], considering its limitation of the process of knowledge gathering and extraction, and being used to develop its Reference View. The first phase, "Knowledge Gathering" (APKE1), consists in gathering knowledge from various sources in IPDP, i.e. Customer Relationship Management (CRM) information, Quality Function Deployment (QFD) information, Computer-Aided Design (CAD) drawings, Computer-Aided Manufacturing (CAM) information, and Computer-Aided Engineering (CAE) simulations.

Sequentially, the "Knowledge Pooling" (APKE2) occurs through a software application that reads the gathered information and extracts its features into an ".xml" extension file that represents the hierarchy of information and their properties. This software is a specialist software, referenced in this paper as "Approach for Product Knowledge Extraction System" (APKE-Sys) that must be developed considering its specific context, the organisation in which the approach is applied.

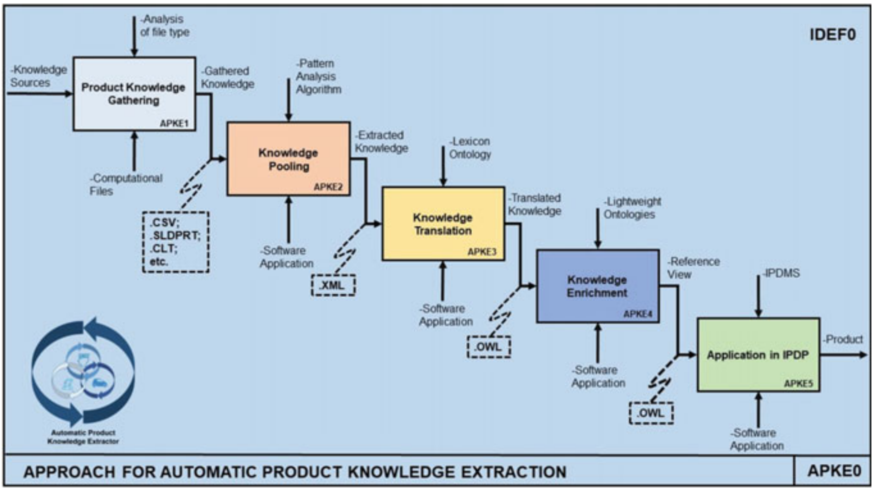

Fig. 2 Proposed approach

Before extracted knowledge can be used "safely", it must be translated and adapted in a way that ensures semantic integrity. This proves that there is still a problem regarding the terminologies used in companies, which may not be in agreement with the literature and/or research. To the translation to occur, a lexicon of the terms from the enterprise is used in the "Knowledge Translation" (APKE3) phase. The lexicon owns enterprise's concepts of products, the context associated with these concepts and their explicit meaning, being formalised in an ontology. The information gathered from the product is translated into information that is useful to aid product development, where the gathered information in ".xml" extension is converted to an ontology format (.owl) and compared to the lexicon ontology through ontology mapping algorithms in the APKE-Sys. The result is an ontology that contains the translated knowledge. The translation process occurs through mapping and intersection using a three-level similarity analysis done by the APKE-Sys:

1. Critical requirement similarities: a comparison between the critical requirements collected from external sources and the concepts in the lexicon ontology;
2. Relationship analysis: a comparison between the relations of concepts in external sources and the relations present in the lexicon ontology;
3. Concept relationship: an analysis of concepts from external sources, regarding their similarity to the lexicon ontology.

In the fourth phase, "Knowledge Enrichment" (APKE4), the ontology that contains the translated knowledge is compared, through ontology mapping in the APKE-Sys, to lightweight ontologies that represent the domains (such as product, design, manufacturing) by a conceptual perspective. Those lightweight ontologies are related to consolidated models that represent their respective knowledge. The results of this phase are the ontologies that compose the Reference View of the IPDMS model, as seen on [23]. Lastly, the "Application in IPDP" (APKE5) phase comprises the addition of semantic rules to the reference ontologies and further application in the IPDP through the IPDMS. In this phase, a consistency analysis of the ontologies is done through an inference engine before and after the creation of the semantic rules, to check for inconsistencies in the mapping processes. The extraction, formalisation, and translation of knowledge to standardised representation can improve the implementation of the IPDMS and, consequently, the IPDP.

6 Experimental Case

6.1 Problem in Industrial Scenario

The application of the proposed approach was carried out in a Brazilian electronics manufacturer, here referred to as Company X. The company had issues related to poor communication in product development, and a few of its products had a high return rate. The company is currently implementing the IPDMS to coordinate its product development and wanted to use a more automated approach to extract product knowledge coming from customer-related data. In this case, which is experimental in nature, the chosen product was a 20 kVA Uninterrupted Power Supply (UPS). This version of the product took around 20 months of development, using more than 2800 h of work and costing around US$ 33.00000 (approximately). During the 20 months of development, the project entered in a 6-month hiatus due to reviews that were necessary, in order to the project be in attendance with needs from customers—this hiatus costs around US$ 13.43200 (approximately) to the enterprise. Furthermore, mistakes in the design caused malfunction while in use, after its launch, bringing more than 70% of products back to the manufacturer.

6.2 The Approach Application

The development process was brought back for its early stages for a full revaluation of its requirements and serves as the case for application of the proposed approach. The application is related to the early stages of PDP, more specifically in the requirements gathering and conversion into product features. The application is outlined in Fig. 3.

Firstly, in the "Knowledge Gathering" phase, the QFD of the new design of the UPS is collected in a **".csv" format** and stored in a folder, accessed by a computational system, here called Approach for Product Knowledge Extraction System (APKE-Sys), that orchestrates the approach application (Detail A of Fig. 3).

In the "Knowledge Pooling" phase, the **".csv" from QFD** was analysed by a **pattern analysis** algorithm in the APKE-Sys, in order to generate its tags and further structure the knowledge from the QFD in an ".xml" extension file (Detail B of Fig. 3). The ".csv" is analysed for keywords related to product, manufacturing, and design parameters. The identified patterns are put in **".xml" tags**, later being

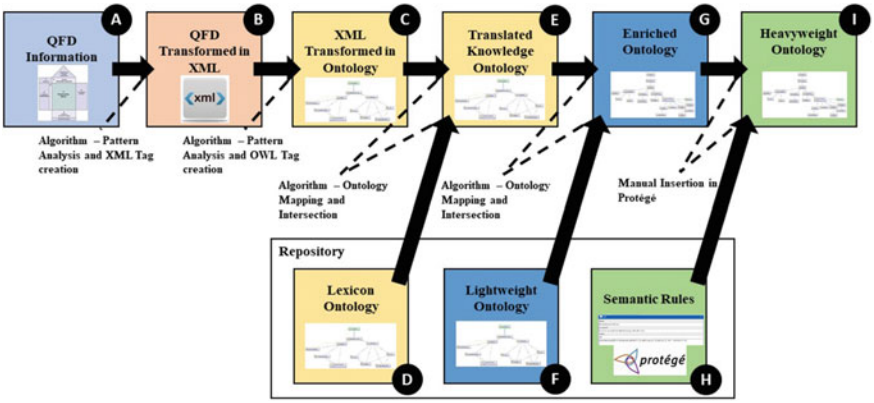

Fig. 3 Application of the proposed approach

joined and forming a file that contains the extracted knowledge from the QFD (Detail B of Fig. 3). This process basically organises the client's requirements to a hierarchical structure (i.e. a "change in the circuit board position" is converted into "Design → UPS → 20 kVA → Circuit Board" in the hierarchical structure).

For the "Knowledge Translation", the tags of the XML file were analysed by the APKE-Sys and **converted** into **".owl" tags**, creating an ontology based on the product knowledge extracted from the QFD (Detail C of Fig. 3). The basic structures of the ".owl" tags are kept in the system and are applied by an algorithm based on **pattern analysis**. Sequentially, the created "QFD Ontology" undergoes an **ontology mapping** process with the lexicon ontology (Detail D of Fig. 3). The commonalities between both ontologies are mapped, and an intersection of the ontology occurs, generating the "Translated Knowledge Ontology" (Detail E of Fig. 3). This phase merely creates an ontological structure of the XML file in accordance with company's specific nomenclatures (i.e. "20 kvA" is renamed as company's code "123ABC").

The "Knowledge Enrichment" process was, as his predecessor, a discrete process in the APKE-Sys, mapping the "Translated Knowledge" Ontology and the chosen domain (product, design, manufacturing) Lightweight Ontology (Detail F of Fig. 3) and combining both in an "Enriched" ontology (Detail G of Fig. 3). Like the process of the "Knowledge Translation", an algorithm for **ontology mapping** looks for similarities in both ontologies (similar classes, attributes and relations), adding the complementary information from the "Lightweight Ontology". This phase gathers the knowledge from the "Translated Knowledge Ontology" and distributes to their specific domain.

The "Application on IPDP" process was performed partially by the APKE-Sys and partially by the IPDMS. The Reference View, an input for application of IPDMS, was generated at the end of the "Knowledge Enrichment" process, by the addition **of semantic rules and consistency analysis in the Protégé software** (Detail H of Fig. 3) that define constraints. The final version of the ontologies offers improved semantics and improved quality on product knowledge. Those are Heavyweight Ontologies (Detail I of Fig. 3) that were validated by a team of specialists, in order to check their overall consistency with company nomenclature.

Firstly, in terms of time saving, the approach reduced an approximate total of 3 h of work from three professionals (9 h total in terms of cost) into a 15 min activity from one single professional (not counting the creation of semantic rules). This is translated into an improvement in time efficiency of more than 97% of previous development, while increasing product's quality with improved communication and semantic correctness, reducing design flaws. This reduction occurs through inconsistencies found on the reasoning process in the Heavyweight Ontology after the application in IPDMS. In terms of operational costs, the final cost of this process was reduced to US$ 292, representing a reduction of more than 97%.

7 Conclusion

The proposed method was able to standardise and extract the product's requirements, reducing the time and the cost of the project without reducing the quality of the final product in many aspects. This standardisation enabled the integration of product and manufacturing and presented reduced misinterpretations in product development.

This research provided a formalisation for the process of capturing information and improved communication and information sharing. The provided method approaches the three main identified issues in the literature review. The consistency analysis of information based on ontology mapping through the phases of Knowledge Translation and Knowledge Enrichment is one of the main strengths of the model, avoiding semantic heterogeneity and human mistakes while structuring the product's requirements.

Next steps of the research will focus in an expansion of the approach, adding more features to different cases, as means to explore the approach and stress its limitations, refining it further.

Acknowledgements The researchers would like to thank the company where the approach was applied, the DAI Program of the Brazilian National Council of Scientific and Technological Development (CNPq), and the Pontifical Catholic University of Paraná, for the funding and structure.

References

1. Szejka, A. L., Júnior, O. C., Loures, E. R., Panetto, H., & Aubry, A. (2016). Proposal of a model-driven ontology for product development process interoperability and information sharing. In *IFIP International Conference on Product Lifecycle Management* (pp. 158–168).
2. Pereira, J. A., & Junior, O. C. (2014). Product development model oriented for the R&D projects of the Brazilian electricity sector. *Applied Mechanics and Materials*, 366–373.
3. Chungoora, N., Young, R. I., Gunendran, G., Palmer, C., Usman, Z., Anjum, N. A., Cutting-Decelle, A. F., Harding, J. A., & Case, K. (2013). A model-driven ontology approach for manufacturing system interoperability and knowledge sharing. *Computers in Industry, 64*(4), 392–401.
4. Andreasen, M. M., & Hein, L. (1987). *Integrated product development*. IFS Publications Ltd.
5. Sosa, M. E., Eppinger, S. D., & Rowles, C. M. (2004). The misalignment of product architecture and organizational structure in complex product development. *Management Science, 50*(12), 1674–1689.
6. Pereira, J. A.. & Junior, O. C. (2014). Product development model oriented for the R&D projects of the Brazilian electricity sector. *Applied Mechanics and Materials, 518*, 366–373.
7. Panetto, H., Dassisti, M., & Tursi, A. (2012). ONTO-PDM: product-driven ontology for product data management interoperability within manufacturing process environment. *Advanced Engineering Informatics, 26*(2), 334–348.
8. Chungoora, N., & Young, R. I. M. (2011). Semantic reconciliation across design and manufacturing knowledge models: A logic-based approach. *Applied Ontology, 6*(4), 295–315.
9. Imran, M., & Young, B. (2013). The application of common logic based formal ontologies to assembly knowledge sharing. *Journal of Intelligent Manufacturing, 26*(1), 139–158.

10. Palmer, C., Urwin, E. N., Pinazo-Sánchez, J. M., Cid, F. S., Rodríguez, E. P., Pajkovska-Goceva, S., & Young, R. I. M. (2016). Reference ontologies to support the development of global production network systems. *Computers in Industry, 77*, 48–60.
11. Fensel, D. (2004). Triple-space computing: Semantic web services based on persistent publication of information. In *Intelligence in communication systems* (pp. 43–53).
12. Fahad, M., Moalla, N., Bouras, A., Qadir, M. A., & Farukh, M. (2010). Disjoint-knowledge analysis and preservation in ontology merging process. In *Software engineering advances (ICSEA)* (pp. 422–428).
13. Junior, O. C., & Young, R. I. M. (2003). Information sharing in multiviewpoint injection moulding design and manufacturing. *International Journal of Production Research, 41*(7), 1565–1586.
14. Haveman, S. P., & Bonnema, G. M. (2013). Requirements for high-level models supporting design space exploration in model-based systems engineering. *Procedia Computer Science, 16*, 293–302.
15. Nattermann, R., & Anderl, R. (2010). Approach for a data-management-system and a proceeding-model for the development of adaptronic systems. In *ASME 2010 International Mechanical Engineering Congress and Exposition* (pp. 379–387).
16. Moneva, H., Hamberg, R., & Punter, T. (2011). A design framework for model-based development of complex systems. In *32nd IEEE Real-Time Systems Symposium 2nd Analytical Virtual Integration of Cyber-Physical Systems Workshop*, Vienna.
17. Junior, O. C., & Young, R. I. M. (2010). Information mapping across injection moulding design and manufacture domains. *International Journal of Production Research, 48*(15), 4437–4462.
18. Stechert, C., & Franke, H. J. (2009). Managing requirements as the core of multi-disciplinary product development. *CIRP Journal of Manufacturing Science and Technology, 1*(3), 153–158.
19. Liao, Y., Lezoche, M., Panetto, H., Boudjlida, N., & Loures, E. R. (2015). Semantic annotation for knowledge explicitation in a product lifecycle management context: A survey. *Computers in Industry, 71*, 24–34.
20. Szejka, A. L., Aubry, A., Panetto, H., Junior, O. C., & Loures, E. R. (2014, October). Towards a conceptual framework for requirements interoperability in complex systems engineering. In *OTM Confederated International Conferences* (pp. 229–240).
21. Szejka, A. L., & Junior, O. C. (2017). The application of reference ontologies for semantic interoperability in an integrated product development process in smart factories. *Procedia Manufacturing, 11*, 1375–1384.
22. Cleland-Huang, J., Chang, C. K., Sethi, G., Javvaji, K., Hu, H., & Xia, J. (2002). Automating speculative queries through event-based requirements traceability. In *Requirements engineering* (pp. 289–296).
23. Szejka, A. L., Leite, A. F. C. S. M., Canciglieri, M. B., & Junior, O. C. (2016). Structuring a foundation basis for semantic interoperability in product development process. In *ISPE TE* (pp. 957–966).